SPECIFICATION

Cache systems are premised on the assumption that the information at the remote source has not changed when the resource copy in the cache is accessed. That is, the resource copy in the cache should not be used in lieu of the resource at the remote source if the resource at the remote source has changed. Thus, in addition to removing copies of resources from the cache memory when it is full, the cache controller in a conventional system also removes copies of resources from the cache memory when it is predicted or determined that the source information has changed, because the copy of the resource in the cache memory is outdated, or "stale". The prediction of whether a resource is likely to have changed is also often used in the selection of which resource copy to remove when space becomes unavailable in the cache memory. For example, an image at a web-site may be expected to change less often than text at a web-site, and thus, a cache controller for caching information downloaded from the Internet may retain downloaded image information for a longer average duration than downloaded text information (BRIDGING PARAGRAPH, page 1, line 22 - page 2, line 4).

The expression "semantic type" as used within this context refers to the different connotative meanings that the information contents of resources can have, as perceived by the user. For example, some information content may be perceived as highly volatile (e.g., being of short-term relevance such as web sites dedicated to the results of sport matches, to specific stock market news or currency exchange rates), other information content may be perceived as rather static (e.g., being of long-term relevance such as glossaries on the Internet). Semantic types that can be expected to contain dynamic information, such as news Web sites and weather

Web sites, need a caching policy wherein the copy in the cache memory is selected for replacement based upon the duration of time that the copy has been in the cache memory. Conversely, semantic types that can be expected to relate to static resources, such as encyclopedic information, glossaries, etc., need a more conservative caching policy, such as least- recently-used (LRU) or least-frequently-used (LFU), that are substantially independent of the time duration that the copy remains in the cache memory. Additionally, some semantic types, such as communicated news messages in popular newsgroups or e-mail messages in e-mail archives may employ a combination of caching policies wherein the copy of the resource, or copies of parts of the resource, are initially identified as dynamically changing, then less dynamic, then static (page 3, lines 9-24).

U.S. PATENT 6,061,763 (RUBIN)

Rubin relates to managing memory in a database management system (Abstract; col.1, lines 9-12).

Rubin addresses specific problems associated with LRU/MRU memory management techniques and memory block segmentation of the buffer cache (col.1, line 35 - col.2, line 23).

Rubin therefore proposes to partition the memory into a plurality of buffer caches, each of which is separately addressable. One buffer cache is set aside as a default buffer cache, while the other buffer caches are reserved for specific data objects meeting certain predefined criteria. Those objects meeting the predefined criteria are stored in reserved buffer caches where they are likely to remain for a relatively long period of time (in comparison to data objects stored in the default buffer caches). A buffer cache

may have a plurality of memory pools, each of which contains multiple storage blocks. The storage blocks in a given memory pool are identically sized, while the storage blocks in one memory pool are sized differently from the storage blocks in another memory pool (Abstract; col.2, line 27 - col.3, line 18).

The "predefined criteria" of the method may be any of many possible limitations specified by a user or system developer (col.2, lines 38-40).

The predefined criteria may require that data objects pertain to particular subject matter (e.g., sales of a widget) (col.2, lines 40-42).

The computer system determines whether a given data object meets the predefined criteria by checking a database catalog listing those objects (preferably tables and indexes) meeting the predefined criteria (col.2, lines 42-46).

All buffer caches--reserved and default--are managed by LRU/MRU rules (col.2, lines 55-56; col. 8, lines 44-49).

The user binds objects meeting predefined criteria (e.g., an object pertains to a particular subject such as employment) to particular buffer caches (col.8, lines 62-65).

The memory of the computer system on which the database management system runs, contains only two buffer caches: one entitled "employees" and another entitled "data 0" (i.e., the default buffer cache). The "employees" buffer cache has been reserved for all data objects stored in either the employee table or an index table for the employee table. No other data objects are bound to the employees buffer cache (col.11, lines 18-26).

Rubin relates to database systems. A database is a specific collection of information organized into interrelated tables of data

and specifications of data objects. The collection of data is organized so that its contents can easily be accessed, managed, and updated. The most prevalent type of database is the relational database, a tabular database in which data is defined so that it can be reorganized and accessed in a number of different ways. A distributed database is one that can be dispersed or replicated among different points in a network. An object-oriented programming database is one that is congruent with the data defined in object classes and subclasses. Databases contain aggregations of data records or files, such as sales transactions, product catalogs and inventories, and customer profiles. Typically, a database manager provides users the capabilities of controlling read/write access, specifying report generation, and analyzing usage (see, e.g., whatis.com and http://searchDatabase.techtarget.com/sDefinition/0,,sid13_gci211895,00.html.

INVENTION

The invention relates, among other things, to a method of processing an information resource. The method comprises receiving a copy of the information resource from a remote source, and caching the copy dependent on a semantic type associated with the information resource.

ARGUMENTS

Applicant respectfully traverses the rejection of the claims under §102(a) as being anticipated by the specification, page 1, line 22 - page 2, line 4. This paragraph recites as an example that an image at a web site may be expected to change less often than

text at the web site. Claim 1 comprises the claim limitation "semantic type of the information resource". The expression "semantic type" has been clarified in the spec (see supra). Applicant respectfully submits that a representation of information (e.g., text, an image) is not the same as the semantic type of the information being represented. Moreover, if the distinction were made between text and images for the purpose of applying the differentiating caching in accordance with the invention, a copy of a web page that contained text and images, e.g., a page from CNN, would get mutilated during the time period over which its components got cached. As explained in the specification by way of example, some resources are preferably cached for a certain time period, others are cached using a LRU/MRU caching policy. It is clearly neither the intention nor the rationale of the invention to cache the text of the CNN page under a policy different from caching an image of the same page.

Accordingly, the rejection of the independent claims under §102(a) is incorrect. The dependent claims recite subject matter patentable over claim 1. For example, claim 3 recites determining a context of a request received from a user, and determining the semantic type based on the request. Therefore, the dependent claims are patentable as well.

Applicant respectfully traverses the rejection of the claims under §102(e) as being anticipated by Rubin.

Rubin uses one buffer cache as a default buffer cache. The other buffer caches are reserved for specific data objects meeting certain predefined criteria. Those specific data objects are likely to remain for a relatively long period of time in the other buffer

caches in comparison to data objects stored in the default buffer caches. The "predefined criteria" may be any of many possible limitations specified by a user or system developer of the database system. The criteria mentioned by Rubin require that the data object in the database pertains to particular subject matter as represented in the pre-indexed database, e.g., "sales of a widget" or "employee name".

Rubin neither teaches nor suggests the caching based on a semantic type of the information resource. Rubin's criterion "pertaining to particular subject matter" neither teaches nor suggests the semantic type in the invention. Applicant respectfully submits that information resources relating to specific subject matter may be of different semantic types. For example, a glossary relating to a certain technical field (subject matter) is rather static (spec., page 3, lines 10-14), whereas press releases from a company about the release of their new product in this same technical field (subject matter) is rather volatile (spec., page 3, lines 14-15). Note that Rubin deals with database management. A database is a specific collection of information organized into interrelated tables of data and specifications of data objects as explained earlier. The semantic attribute of the invention as claimed would not appear in the database environment of Rubin.

Accordingly, the rejection of the independent claims under 102(e) as anticipated by Rubin is incorrect. The dependent claims recite subject matter patentable over the independent claims and are therefore allowable as well.

Applicant respectfully submits that the rejection of dependent under 103(a) is incorrect in view of what has been explained above

with respect to Rubin.

Applicant respectfully submits that she has answered all issues raised by the Examiner and that the application is in condition for allowance. Such allowance is respectfully requested.

Please charge any fees other than the issue fee to deposit account 14-1270. Please credit any overpayments to the same account.

CERTIFICATE OF MAILING
I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to:
Commissioner of Patents and Trademarks
Washington, D.C. 20231
On (date): July 30, 2001
By: [signature]

Respectfully submitted,

By [signature]

Peter Verdonk
Limited Recognition under 37 CFR 10.9(b)

Tel. (408) 617-4613

CLEAN COPY OF PAGE 1 WITH AMENDED TITLE IN 09/374,694 (attorney docket PHA 23,737)

SEMANTICS-BASED CACHING POLICY TO MINIMIZE LATENCY

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates to the field of information processing systems, and in particular to information processing systems that utilize cache memory to minimize latency.

2. Description of Related Art

Cache systems are common in the art. A cache system comprises a cache memory and a corresponding controller that regulates the storage and retrieval of information to and from the cache memory. Traditionally, the cache memory is filled with copies of information resources that a user receives from a remote source, "remote" being defined as being further removed from the user than the cache memory, e.g., local main memory or a server in a client-server architecture. If the user subsequently requests the same resource, the resource's copy is provided from the cache memory, rather than from the original remote source, thereby saving the time required to receive the resource from the remote source for a second time. When the cache memory becomes full, the cache controller removes copies of the resources that have not been accessed recently, to make room for copies of new resources that the user accesses. A variety of criteria, commonly termed caching policies, are available to determine which resource copy to remove from the cache memory. Such caching policies can be based on: the duration since the last access, the number of times accessed since originally received, the amount of memory allocated to the resource, the difficulty of retrieving the resource from the remote site, etc.

Cache systems are premised on the assumption that the information at the remote source has not changed when the resource copy in the cache is accessed. That is, the resource copy in the cache should not be used in lieu of the resource at the remote source if the resource at the remote source has changed. Thus, in addition to removing copies of resources from the cache memory when it is full, the cache controller in a conventional system also removes copies of resources from the cache memory when it is predicted or determined that the source information has changed, because the copy of the resource in the cache memory is outdated, or "stale". The prediction of whether a resource is likely to have changed is also often used in the selection of